

## Supporting Information: Probing Interactions Between Histone Tails and Nucleosomal DNA via Product and Kinetic Analysis.

Liwei Weng, Chuanzheng Zhou, and Marc M. Greenberg\*

Department of Chemistry  
Johns Hopkins University  
3400 N. Charles St.  
Baltimore, MD 21218

[mgreenberg@jhu.edu](mailto:mgreenberg@jhu.edu); phone: 410-516-8095

### Contents:

1. General Methods. (S1-7)
2. **Figure S1.** Oligonucleotides synthesized for preparing 145 nt single stranded 601 DNA. (S8-9)
3. **Figure S2.** Preparation of 145mer ssDNA by ligation. (S10)
4. **Figure S3.** Sequence of 601 DNA and positions of modifications. (S10)
5. **Figure S4.** Preparation of internally radiolabeled strands. (S11)
6. **Figure S5.** Representative SDS and TAU PAGE identifying the cross-linked proteins with **4**. (S12)
7. **Figure S6.** Kinetic study of **4**<sub>89</sub> in NCPs containing H4 variants. (S13-15)
8. **Figure S7.** Structures of protected amino acid (**6a-i**). (S16)
9. **Figure S8.** Structures of amino acid adducts (**7a-d**) from monomer reactions with **4**. (S16)
10. **Figure S9.** Chromatograms of **4** reacting with amino acids (**6a-c**). (S17-18)
11. **Figure S10.** Representative chromatograms of purifying oligonucleotides containing **3**. (S19)
12. **Figure S11-S13.** DNase I footprinting on NCPs with **3** and H4 variants. (S20-22)
13. **Figure S14.** ESI-MS spectra of modified oligonucleotides. (S23-24)
14. **Figure S15.** MALDI-TOF mass spectra of modified oligonucleotides. (S25-29)
15. **Figure S16.** ESI mass spectra of histone H4 K8, 12, 16, 20R and K8, 12,16, 20R/H18A variants. (S30)
16. **Figure S17.** LC-MS/MS spectra of peptide fragments following thermolysin digestion of histone H4. (S31)
17. **Table S1.** Calculated and observed fragmentation patterns for peptides in Figures 2, 3, and Figure S17. (S32)

**General Methods.** Oligonucleotides containing phenyl selenide or C4-AP precursors (**3** and **2**) were synthesized and purified as previously described.<sup>[1,2]</sup> DNA containing **3** was subjected to additional purification by reversed-phase HPLC on a RP-C<sub>18</sub> column (VARIAN, Microsorb-MV 100-5 C<sub>18</sub> 250 4.6 mm). The peak of interest was collected using the following gradient conditions: 0-5 min 0-2% B in A, 5-15 min 2-12% B in A, 15-40 min 12-20% B in A, 40-50 min 20-30% B in A, 50-55 min 30-80% B in A, 55-60 min 100% A, at a flow rate 1.0 mL/min. [A: 0.05 M (Et<sub>3</sub>NH)OAc (pH 7.0)/MeCN 95:5; B: 0.05 M (Et<sub>3</sub>NH)OAc (pH 7.0)/MeCN 50:50]. Expression and purification of core histone

proteins, as well as refolding and purification of the histone octamer, were carried out as previously described.<sup>[3]</sup> T4 polynucleotide kinase, T4 DNA ligase, and DNase I were purchased from New England Biolabs (NEB). Nuclease P1 (from *Penicillium citrinum*) was from Sigma and was dissolved in water (1 U/ $\mu$ L). Benzonase was purchased from Sigma.  $\gamma$ -<sup>32</sup>P-ATP was purchased from Perkin Elmer. C<sub>18</sub>-Sep-Pak cartridges were obtained from Waters. Quantification of radiolabeled oligonucleotides was carried out using a Molecular Dynamics phosphorimager equipped with ImageQuant Version TL software. Salmon sperm DNA (10 mg/mL) was purchased from Invitrogen. All experiments with NCPs were conducted in clear siliconized tubes (Bio Plas Incorporated). Photolyses of oligonucleotides were carried out in a Rayonet photoreactor (RPR-100) equipped with 16 lamps with a maximum output at 350 nm. ESI-MS analysis was carried out on a LCQ-Deca Ion Trap. UPLC/MS system is an Acquity UPLC H-Class/Xevo G2 QToF from Waters. MALDI-TOF MS analysis was carried out on Bruker AutoFlexIII MALDI-TOF.

**Thermolysin digestion of isolated H4.** The preparation of NCPs containing **2** in large scale was described previously.<sup>[4]</sup> After incubating NCP (~ 4 nmol) containing C4-AP<sub>89</sub> and C4-AP<sub>234</sub> at 37 °C for 8 h, 1  $\times$  benzonase buffer (50 mM Tris•HCl, pH 8.8, 10 mM MgCl<sub>2</sub>, 20 mM NaCl) was added, followed by the addition of 1  $\mu$ L (250 units) of benzonase. The four histone proteins were separated by RP-HPLC as previously described.<sup>[4]</sup> HPLC purified H4 (~ 1 nmol, 10  $\mu$ g) and thermolysin (0.5  $\mu$ g) were mixed in a 50  $\mu$ L solution containing 50 mM NH<sub>4</sub>HCO<sub>3</sub> and 1 mM CaCl<sub>2</sub>. The reaction mixture was incubated at 37 °C for 3 h, quenched with TFA to a final concentration of 1%. An aliquot (10  $\mu$ L) was analyzed by UPLC-MS/MS. The column used was an ACQUITY UPLC HSS T3 Column, 100Å, 1.8  $\mu$ m, 2.1 mm  $\times$  100 mm (From Waters). Buffer A is water, Buffer B is MeCN, Buffer C is 1% formic acid. Gradient: 0-1-36 min, B from 5%-5%-40%, keep C at 10% constantly. Column temperature: 35 °C, flow rate: 0.3 mL / min.

**In-gel acetylation and trypsin digestion of H4.** The histone H4 with modification was purified and subjected to in-gel acetylation followed by trypsin digestion as described previously.<sup>[4]</sup> After drying the digested peptides in speedvac, the residue was dissolved in 0.1% TFA (30  $\mu$ L). An aliquot (10  $\mu$ L) was subjected to UPLC-MS/MS analysis under the same conditions as described above.

**General procedure for the oxidation reaction of NCPs containing 3.** The ligation of 145 mer 601 DNA containing **3** and the reconstitution of NCPs were carried out as described previously.<sup>[5]</sup> However, the oligonucleotides and ligated strands were different (Figure S2). To the reconstituted NCP solution was added 5 mM NaIO<sub>4</sub>. Following incubation at 37 °C for 2 h, the samples were divided into two portions. One portion was treated with proteinase K (0.1  $\mu$ g) for 5 min at room temperature and analyzed by 8% denaturing PAGE (40  $\times$  32  $\times$  0.04 cm). To the second portion was directly added 4  $\times$  SDS loading buffer (400 mM Tris•HCl, 400 mM DTT, 8% SDS, 40% glycerol) and analyzed by SDS PAGE (10% resolving acrylamide/bisacrylamide = 29:1, 5% stacking layer, 20  $\times$  16  $\times$  0.1 cm). The gel was run at 250 V until the bromophenol blue band migrated to the bottom.

**General procedure for the kinetic studies.** NCPs containing 5'-radiolabeled DNA duplex were mixed with 3 mM NaIO<sub>4</sub> without additional buffer. Na<sub>2</sub>SO<sub>3</sub> (30 mM) was added 5 min afterwards to react with excess NaIO<sub>4</sub>. An aliquot was withdrawn from the reaction and immediately put in -80 °C at each indicated time point. The aliquot was mixed with 4  $\times$  SDS loading buffer before being subjected to SDS PAGE analysis (10% resolving layer, acrylamide/bisacrylamide 29:1, 5% stacking layer, 20  $\times$  16  $\times$  0.1 cm).

**LC-MS analysis for the reactions between phenyl-selenide-containing monomer (3) and protected amino acids.** Typically, monomer **3** (1 mM) was mixed with protected amino acid **6a-f** (50 mM) and 10 mM sodium phosphate buffer (pH 7.2), to which NaIO<sub>4</sub> (5 mM) was added. If the removal of excess NaIO<sub>4</sub> was required, Na<sub>2</sub>SO<sub>3</sub> (50 mM) was added to the mixture 5 min after the

addition of NaIO<sub>4</sub>. The resulting solution was incubated at 37 °C for overnight. The reaction was diluted by 5-fold prior to being subjected to UPLC-MS analysis. The column and the temperature of the column was the same as mentioned above. The separation condition is 0.1 % formic acid in water (solvent A) and acetonitrile (solvent B), 0.3 mL/min and using the following linear gradient: (time (min), % B) 5, 0; 18, 10; 25, 25; 30, 90; 35, 0. Mass spectra were acquired in positive ion mode with MSE using a capillary voltage of 3 kV, a sample cone voltage of 30 V and an extraction cone voltage of 4 V. The cone gas flow was set up to 30 L/h and desolvation gas flow was 800 L/h. Desolvation temperature and source temperature were set to 400 and 150 °C, respectively. The acquisition range was m/z 100-3000. The scan acquisition rate was 10 Hz. The LC-MS system was operated by the Mass Lynx software.

**DNase I footprinting of NCPs.** To the concentrated NCP solution (from 240 µL to ~ 40 µL by 10K Amicon concentration device) was added 2 µL of 10 × DNase I buffer (10 mM Tris•HCl, 2.5 mM MgCl<sub>2</sub>, 0.5 mM CaCl<sub>2</sub>, pH 7.6) and indicated amounts of DNase I (Figure S11-S12). The reaction was carried out for 5 min at room temperature and quenched by adding 1 µL of 0.5 M EDTA. Nucleoprotein gel electrophoresis (10 × 8 × 0.15 cm, 6% acrylamide/bisacrylamide, 59:1, 0.6 × TBE buffer) was used to purify NCP with cleaved DNA. The intact NCP band was excised from the gel and the DNA was eluted overnight in 500 µL of elution buffer (0.2 M NaCl and 1 mM EDTA) containing 0.1% SDS. The gel particle was filtered using a polyprep-column (BioRad) and the filtrate was mixed with 1 µg of salmon sperm DNA followed by ethanol precipitation. The obtained samples were analyzed by 8% denaturing PAGE (40 × 32 × 0.04 cm).

**Determining the protein(s) involved in cross-linking with 4.** DNA duplexes containing 5'-<sup>32</sup>P-3 were assembled in the same manner as previously described<sup>[4]</sup> except the oligonucleotides and ligated duplexes were those depicted in Figure S4. To the NCP solution (typically 240 µL) was added 5 mM

NaIO<sub>4</sub>. Following incubation for 4 h, the mixture was concentrated with Amicon Ultra 10K centrifugal filter to 50  $\mu$ L. The resulting solution was diluted to 500  $\mu$ L of 0.1% SDS in water followed by concentration again to 50  $\mu$ L in the same Amicon tube. An aqueous solution of benzonase (2 U, 0.5  $\mu$ L) and buffer are added to create a 100  $\mu$ L solution containing 50 mM Tris•HCl pH 8.0, 2 mM MgCl<sub>2</sub>, 20 mM NaCl. After incubation at 37 °C for 2 h, the reaction mixture was diluted with 1  $\times$  Antarctic phosphatase buffer (50 mM bis-Tris-Propane•HCl, pH 6.0, 1 mM MgCl<sub>2</sub>, 0.1 mM ZnCl<sub>2</sub>) to 500  $\mu$ L and concentrated using Amicon Ultra 10K centrifugal filter to 50  $\mu$ L. Repeat the dilution-concentration once more in the same Amicon tube. The final volume was adjusted to 100  $\mu$ L, to which was added 2 U nuclease P1. Following 4 h incubation, 900  $\mu$ L of cold acetone was added and the mixture was kept at -20 °C overnight. The proteins were pelleted by centrifugation (16,000 g) and washed with 500  $\mu$ L of cold acetone. After drying, the proteins were analyzed by SDS-PAGE (18% resolving layer, acrylamide/bisacrylamide 29:1, 20  $\times$  16  $\times$  0.1 cm, 5% stacking layer). Samples analyzed by TAU PAGE were dissolved in 8 M urea solution containing 10 mM DTT and 5% acetic acid. TAU gels (15 %, acrylamide/bisacrylamide 59:1, 8 M urea, 5 % acetic acid, 0.37 % Triton X-100) were prepared and eletrophoresed as described previously.<sup>[6]</sup> Wild-type histone octamer were loaded as reference. Both SDS and TAU PAGE were first stained with Coomassie blue, and then subjected to phosphorimaging analysis.

**Preparation of the plasmids for expression of mutated histones.** The plasmids used for expression of mutated H4 were prepared by site-directed mutagenesis. The WT histone H4 (pET3a-H4-WT) plasmid was obtained as a gift from Professor Greg Bowman (JHU Biophysics). Plasmid for H4 K8, 12, 16, 20R was obtained from template of pET3a-H4-K5, 8, 12, 16, 20R, which was previously prepared.<sup>[7]</sup> Similarly, plasmid for H4 K8, 12, 16, 20R, H18A was prepared from template of pET3a-H4-K5, 8, 12, 16, 20R, H18A. The vector is pET3a and inset DNA sequences are listed below.

### Sequence of pET3a-H4-K5, 8, 12, 16, 18R:

-CTTTAAGAAGGAGATACA TATG (start codon)

1

tct ggt cgt ggt cgt ggt ggt cgt ggt ctg ggt cgt ggt ggt gct cgt cgt cac cgt

20

cgt gtt ctg cgt gac aac atc cag ggt atc acc aag ccg gct atc cgt cgt atg gct

39

cgt cgt ggt ggt gtt aaa cgt atc tcc ggt ctg atc tac gaa gaa acc cgc ggt gtt

58

ctg aaa gtt ttc ctg gaa aac gtt atc cgt gac gct gtt acc tac acc gaa cac gct

77

aaa cgt aaa acc gtt acc gct atg gac gtt gtt tac gct ctg aaa cgt cag ggt cgt

96

102

acc ctg tac ggt ttc ggt ggt TAA (stop codon) AGA TCCGGCTGC-

### Sequence of pET3a-H4-K5, 8, 12, 16, 18R, H18A:

-CTTTAAGAAGGAGATACA TATG (start codon)

1

tct ggt cgt ggt cgt ggt ggt cgt ggt ctg ggt cgt ggt ggt gct cgt cgt gct cgt

20

cgt gtt ctg cgt gac aac atc cag ggt atc acc aag ccg gct atc cgt cgt atg gct

39

cgt cgt ggt ggt gtt aaa cgt atc tcc ggt ctg atc tac gaa gaa acc cgc ggt gtt

58

ctg aaa gtt ttc ctg gaa aac gtt atc cgt gac gct gtt acc tac acc gaa cac gct

77

aaa cgt aaa acc gtt acc gct atg gac gtt gtt tac gct ctg aaa cgt cag ggt cgt

96

102

acc ctg tac ggt ttc ggt ggt TAA (stop codon) AGA TCCGGCTGC-

### **Preparation of plasmid for H4 K8, 12, 16, 20R (pET3a-H4-K8, 12, 16, 20R)**

Template: pET3a-H4-K5, 8, 12, 16, 20R

Forward primer: 5'-CATATGTCCTGGTCGTGGTAAAGGTGGTCGTGGTCTGGGTC-3'

Reverse primer: 5'-GACCCAGACCACGACCACCTTTACCACGACCAGACATATG-3'

### **Preparation of plasmid for H4 K8, 12, 16, 20R, H18A (pET3a-H4-K8, 12, 16, 20R, H18A)**

Template: pET3a-H4-K5, 8, 12, 16, 20R, H18A

Forward primer: 5'-CATATGTCCTGGTCGTGGTAAAGGTGGTCGTGGTCTGGGTC-3'

Reverse primer: 5'-GACCCAGACCACGACCACCTTTACCACGACCAGACATATG-3'

[1] Szczepanski, J. T. et al. *J. Am. Chem. Soc.* **2008**, *130*, 9646.

[2] Peng, X. et al. *J. Am. Chem. Soc.* **2008**, *130*, 12890.

[3] Dyer, P. N. et al. *Methods Enzymol.* **2004**, *375*, 23.

[4] Zhou, C. et al. *J. Am. Chem. Soc.* **2013**, *135*, 5274.

[5] Szczepanski, J. T. et al. *Proc. Nat. Acad. Sci. USA* **2010**, *107*, 22475.

[6] Lennox, R. W. et al. *Methods Enzymol.*, **1989**, *170*, 532.

[7] Szczepanski, J. T. *Biochemistry*, **2013**, *52*, 2157

(S1) ATC GAT GTA TAT ATC TGA CAC GTG CCT GGA

(S2) GAC TAG GGA GTA ATC CCC TTG GCG GTT AAA ACG CG

(S3) GGG GAC AGC GCG TAC

(S4) GGG GAC A<sub>3<sub>73</sub></sub>C GCG TAC G

(S5) GTG CGT TT<sub>3<sub>89</sub></sub> AGC GGT GCT AG

(S6) TG CGT TTA AGC GGT GCT AG

(S7) AGC TGT CTA CGA CCA ATT GAG CGG CCT CGG CAC CGG GAT TCT GAT

(S8) CTC CCT AGT CTC CAG GCA CG

(S9) GTA GAC AGC TCT AGC ACC GCT AAA ACG CAC GTA CGC GCT GTC CCC CGC GTT TT

(S10) GGG GAC AGC GCG TAC GTG CGT TTA AGC GGT GCT AG

(S11) AG CTG TCT ACG ACC AAT T<sub>3<sub>119</sub></sub>A GCG GCC T

(S12) CGG CAC CGG GAT TCT GAT

(S13) CGC TGT CCC CCG CGT TTT AA

(S14) GTA GAC AGC TCT AGC ACC GC

(S15) CGG TGC CGA GGC CGC T

(S16) ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TC

(S17) GTA GAC AGC TCT AGC ACC GCT <sub>3<sub>202</sub></sub>AA ACG CAC

(S18) GTA GAC AGC TCT AGC ACC GCT AAA ACG CA

(S19) GTA CGC GCT GTC CCC CGC GTT TT

(S20) CG TAC GCG <sub>3<sub>218</sub></sub>TG TCC CCC GCG TTT TAA

(S21) CCG CCA AGG GGA TTA CTC CCT AGT CTC CAG GCA CGT GTC AGA TAT ATA CAT CGA T

(S22) AGC TGT CTA CGA CCA ATT GA

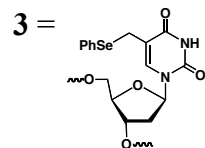
(S23) CGC GTA CGA AAA CGC G

(S24) CCC TTG GCG GTT AAA ACG CG

(S25) GTA GAC AGC TCT AGC ACC GCT AAA A

(S26) <sub>3<sub>205</sub></sub>CG CAC GTA CGC GCT GTC CCC CGC GTT TTA A

(S27) ATC AGA ATC CCG GTG CCG



(S28) AGG CCG CT<sub>3</sub><sub>172</sub> AAT TGG TC

(S29) GGG GAC AGC GCG TAC GTG CGT TTA

(S30) GGG GAC A

(S31) <sub>3</sub><sub>89</sub>AG CGG TGC TAG

(S32) <sub>3</sub><sub>73</sub>CG CGT ACG TGC GTT TAA GCG GTG CTA G

(S33) AGC TGT CTA CGA CCA ATT

(S34) <sub>3</sub><sub>119</sub>AG CGG CCT CGG CAC CGG GAT TCT GAT

(S35) GTA GAC AGC TCT AGC ACC GCT

(S36) GTA GAC AGC TCT AGC ACC GCT AAA ACG CAC GTA CGC G

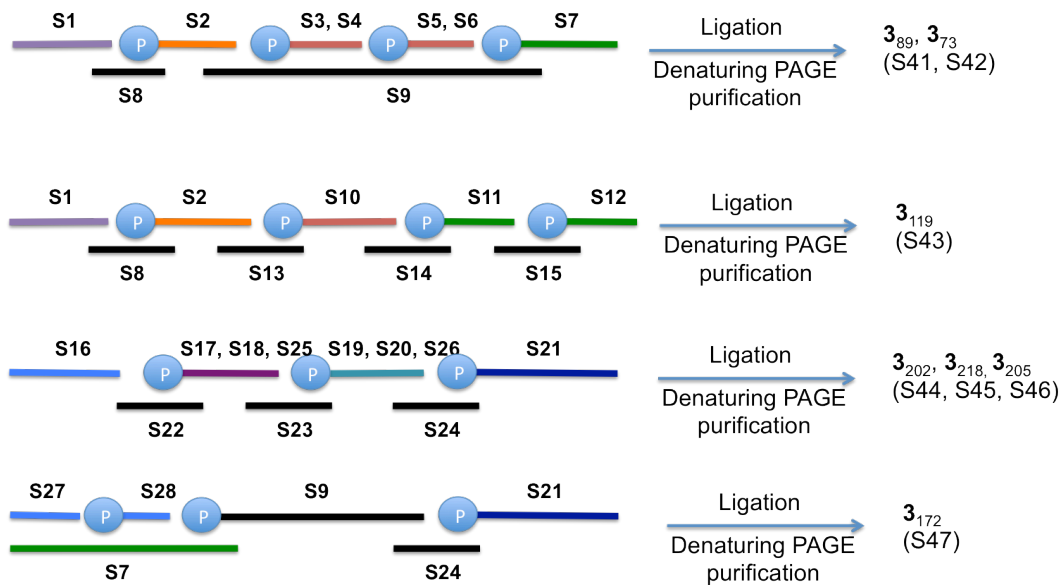
(S37) <sub>3</sub><sub>202</sub>AA ACG CAC GTA CG

(S38) <sub>3</sub><sub>218</sub>TG TCC CCC GCG TTT TAA

(S39) ATC AGA ATC CCG GTG CCG AGG CCG CT

(S40) <sub>3</sub><sub>172</sub>AA TTG GTC

**Figure S1.** Oligonucleotides synthesized for preparing 145 nt single stranded 601 DNA.

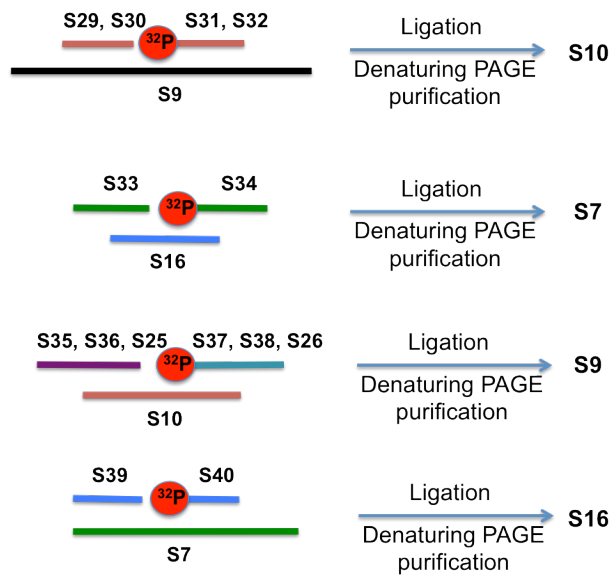


**Figure S2.** Preparation of 145mer ssDNA by ligation.

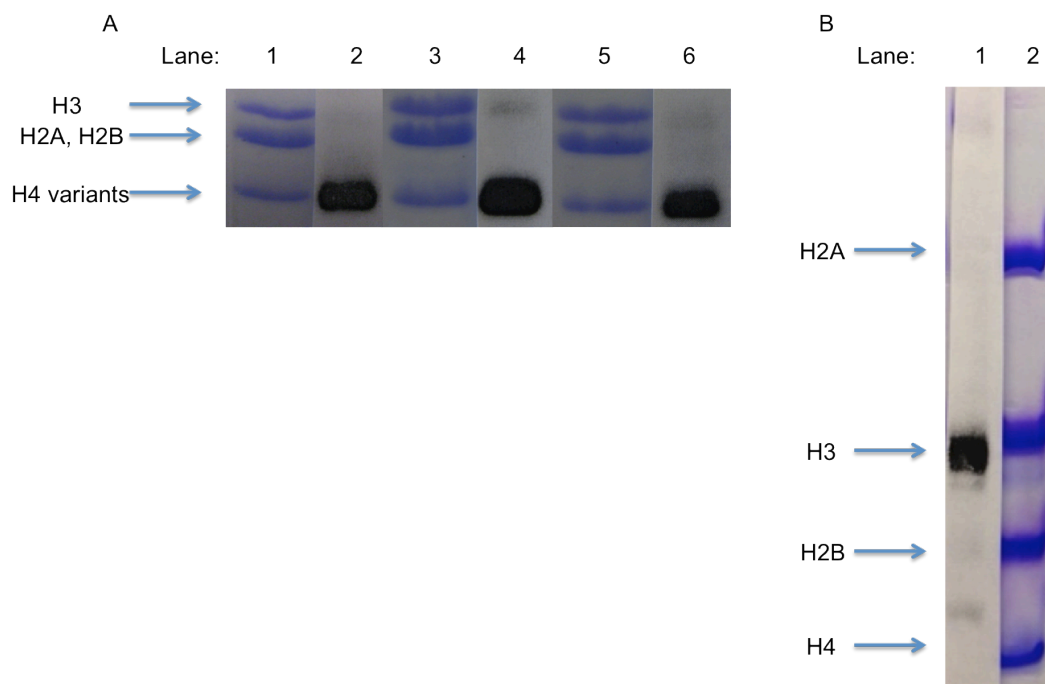
5' -ATCGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCCTTGGCGGTAAAACGCGGGGACAN<sub>73</sub>  
3' -TAGCTACATATATAGACTGTGCACGGACCTCTGATCCCTCATTAGGGGAACCGCCAATTTGCGCCCCCTGTN<sub>218</sub>  
CGCGTACGTGCGT TTN<sub>89</sub> AGCGGTGCTAGAGCTGTCTACGACCAATTN<sub>119</sub> AGCGGCCTCGGCACCGGATTCTGAT-3'  
GCGCATGCACGCA<sub>205</sub> AAN<sub>202</sub> TCGCCACGATCTCGACAGATGCTGGTTAAN<sub>172</sub> TCGCCGGAGCCGTGGCCCTAAGACTA-5'

Substrate	Position of modification	N <sub>73</sub> (AXC)	N <sub>89</sub> (AXC)	N <sub>119</sub> (TXA)	N <sub>172</sub> (TXA)	N <sub>202</sub> (TXA)	N <sub>205</sub> (CXA)	N <sub>218</sub> (TXA)	A <sub>72</sub>	C <sub>74</sub>
S41	SHL 1.5 (89)	G	3	G	C	A	A	C	A	C
S44	SHL 1.5 (202)	G	A	G	C	3	A	C	A	C
S43	SHL 4.5 (119)	G	A	3	A	C	A	C	A	C
S47	SHL 4.5 (172)	G	A	A	3	C	A	C	A	C
S42	SHL dyad (73)	3	A	G	C	C	A	A	A	C
S45	SHL dyad (218)	A	A	G	C	C	A	3	A	C
S48	SHL dyad (73)	3	A	G	C	C	A	A	T	A
S46	SHL 1.2 (205)	G	A	G	C	C	3	C	A	C

**Figure S3.** Sequence of 601 DNA and position of 3.

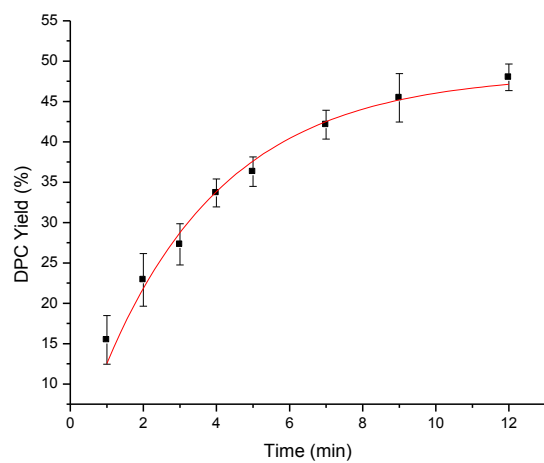


**Figure S4.** Preparation of internally radiolabeled strands.

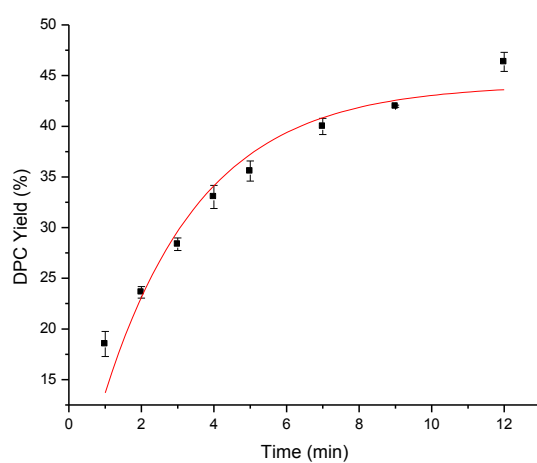


**Figure S5.** Representative SDS and TAU gel image for identifying cross-linked proteins with **4** in NCPs. (A) Example SDS PAGE (18%) of histone proteins following the 5'-<sup>32</sup>P-**3**<sub>89</sub> transfer assay. Lanes 1, 3, 5 are histone octamers with WT H4 (lane 1), H4 H18A (lane 3) and H4 K5, 8, 12, 16, 20R (lane 5) stained by coomassie blue; lanes 2, 4, 6 are the autoradiograms of histone octamers comprising WT H4 (lane 2), H4 H18A (lane 4) and H4 K5, 8, 12, 16, 20R (lane 6) following DPC formation with the 5'-<sup>32</sup>P-**3**<sub>89</sub> and subsequent DNA digestion. (B) Sample TAU PAGE (20%) of histone proteins following 5'-<sup>32</sup>P-**3**<sub>205</sub> transfer assay. Lane 1 is coomassie blue stained WT histone octamer. Lane 2 is the autoradiogram of the histone octamers composed of WT histone proteins following DPC formation and subsequent DNA digestion.

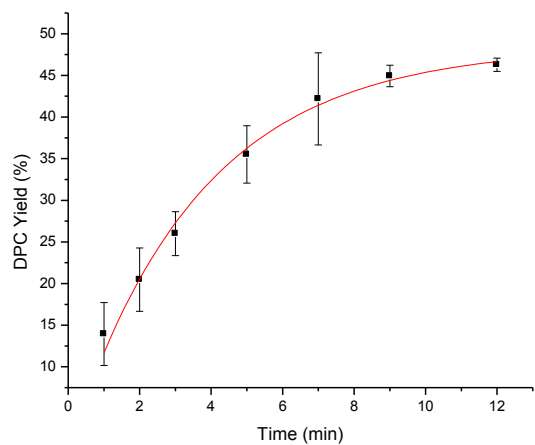
(A) K16A



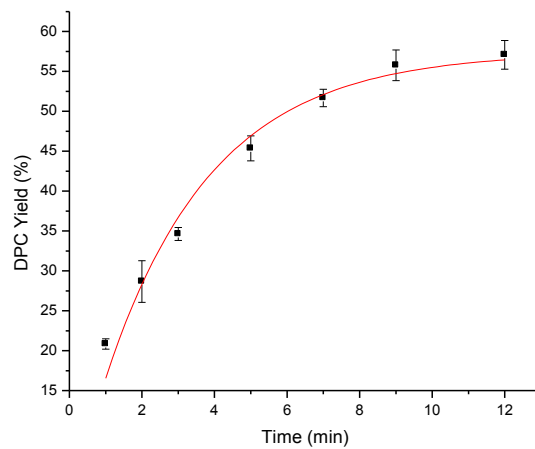
(B) K20A



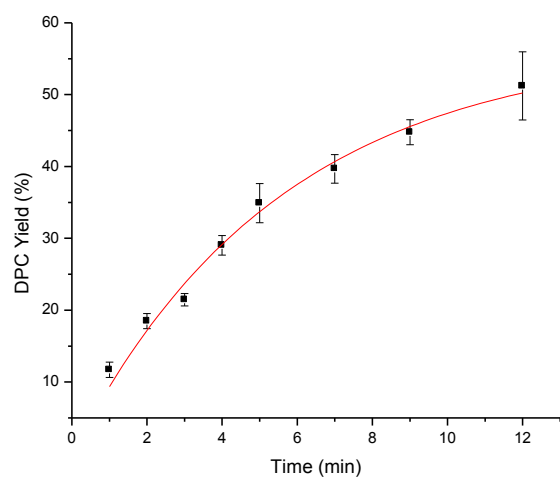
(C) K16, 20A



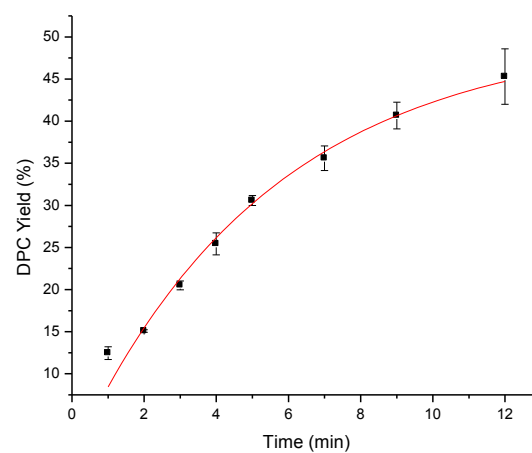
(D) K5, 8, 12A



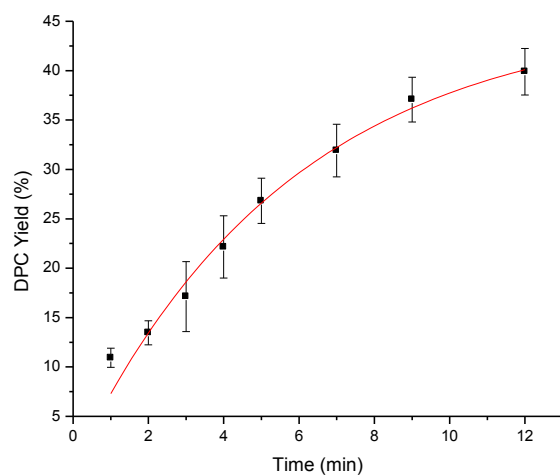
(E) H18A



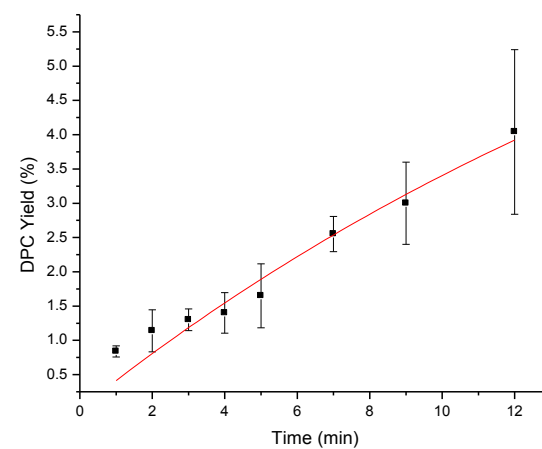
(F) K5, 8, 12, 16, 20R



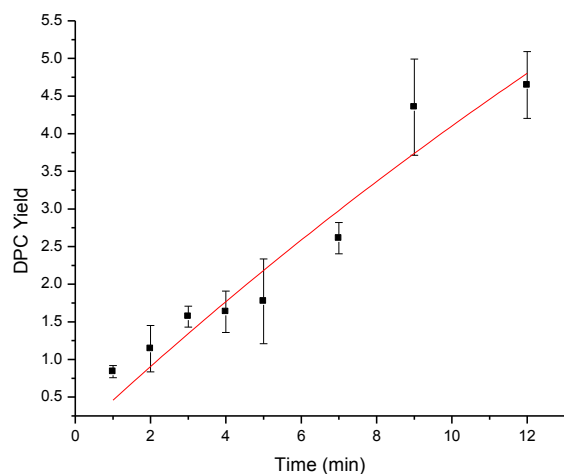
(G) K8, 12, 16, 20R



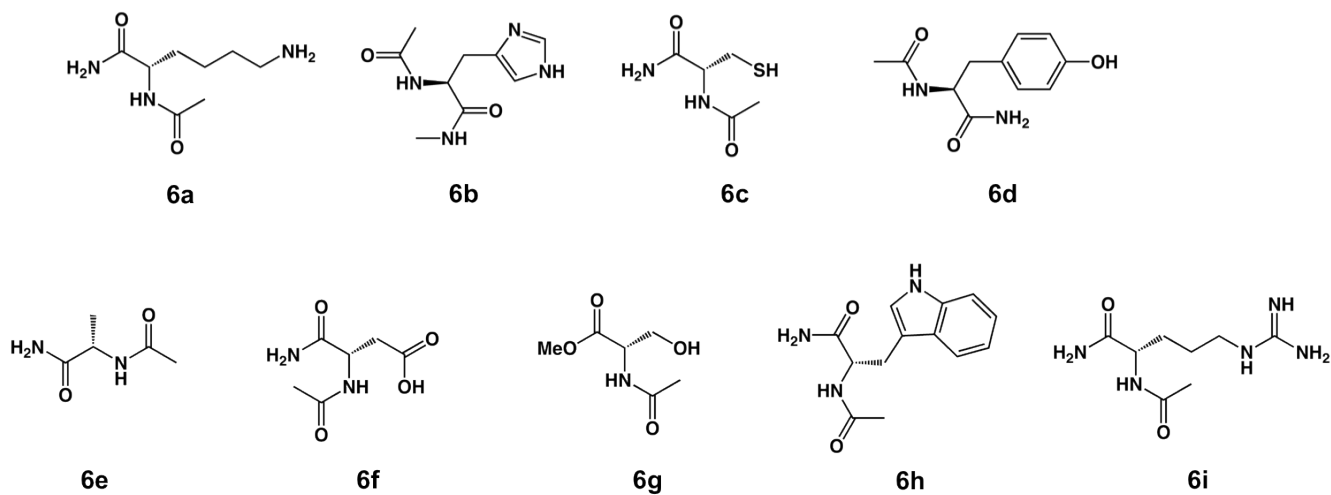
(H) K5, 8, 12, 16, 20R/H18A



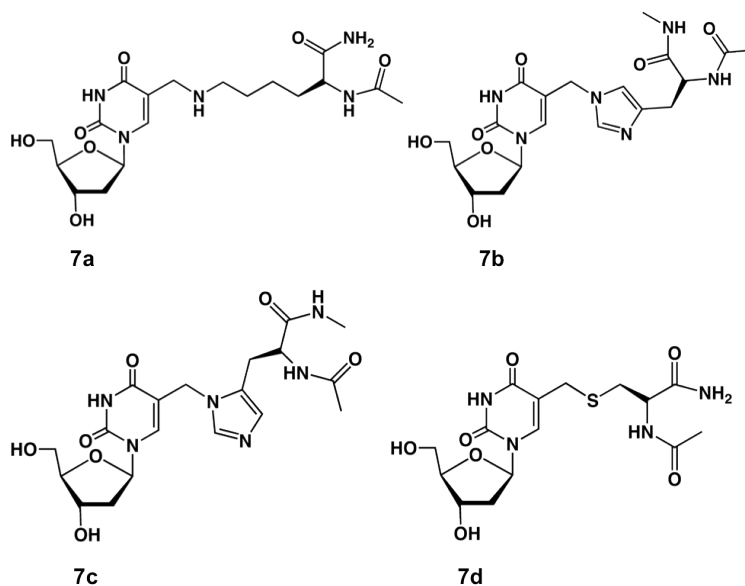
(I)



**Figure S6.** Kinetic study of  $4_{89}$  in NCPs containing (A) H4 K16A variant; (B) H4 K20A variant; (C) H4 K16,20A variant; (D) H4 K5, 8, 12A variant; (E) H4 H18A variant; (F) H4 K5, 8, 12, 16, 20R variant; (G) H4 K8, 12, 16, 20R variant; (H) H4 K5, 8, 12, 16, 20R/H18A variant; (I) H4 K8, 12, 16, 20R/H18A.

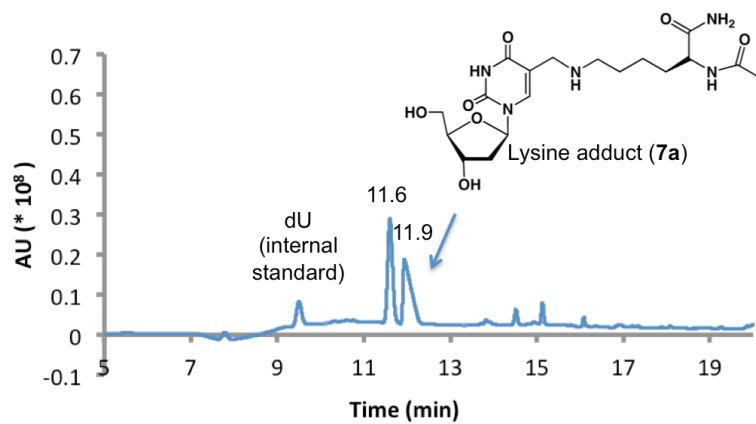


**Figure S7.** Structures of protected amino acids used in the monomer study: lysine (**6a**), histidine (**6b**), cysteine (**6c**), tyrosine (**6d**), alanine (**6e**), aspartic acid (**6f**), serine (**6g**), tryptophan (**6h**) and arginine (**6i**).

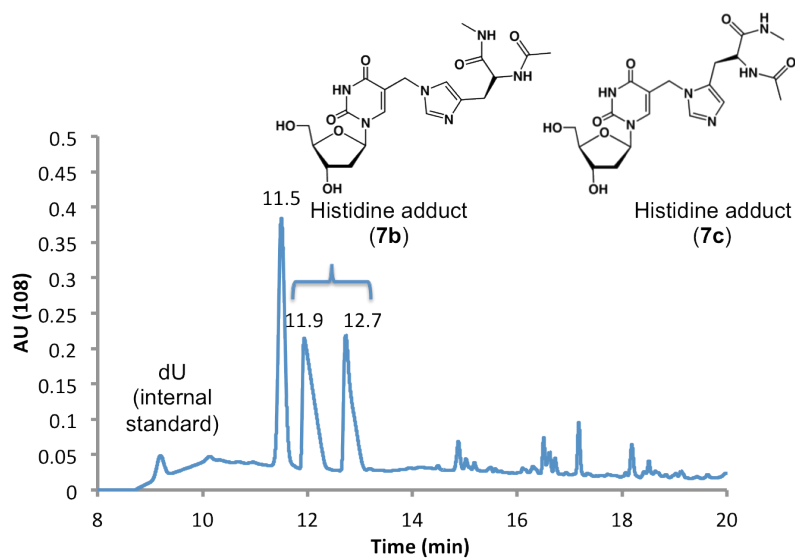


**Figure S8.** Structures of amino acid adducts from monomer reactions with **4**: lysine adduct (**7a**), histidine adducts (**7b** and **7c**), cysteine adduct (**7d**).

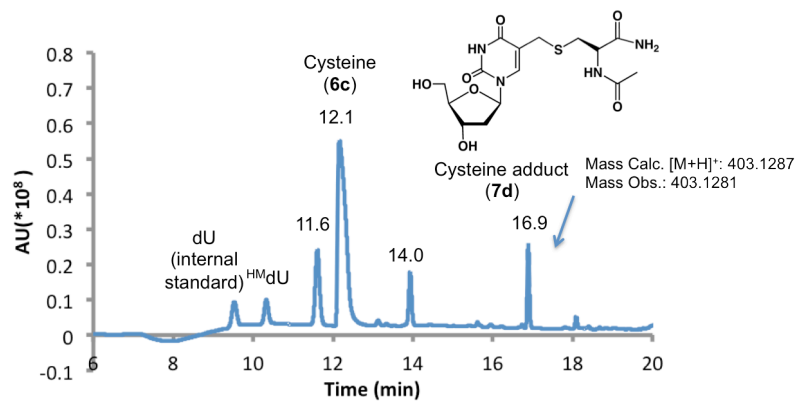
(A)



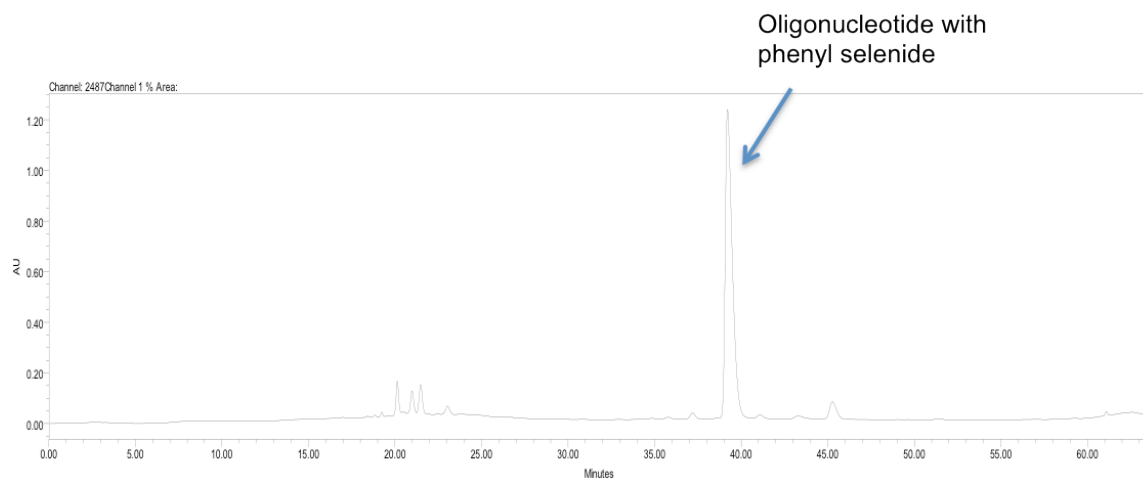
(B)



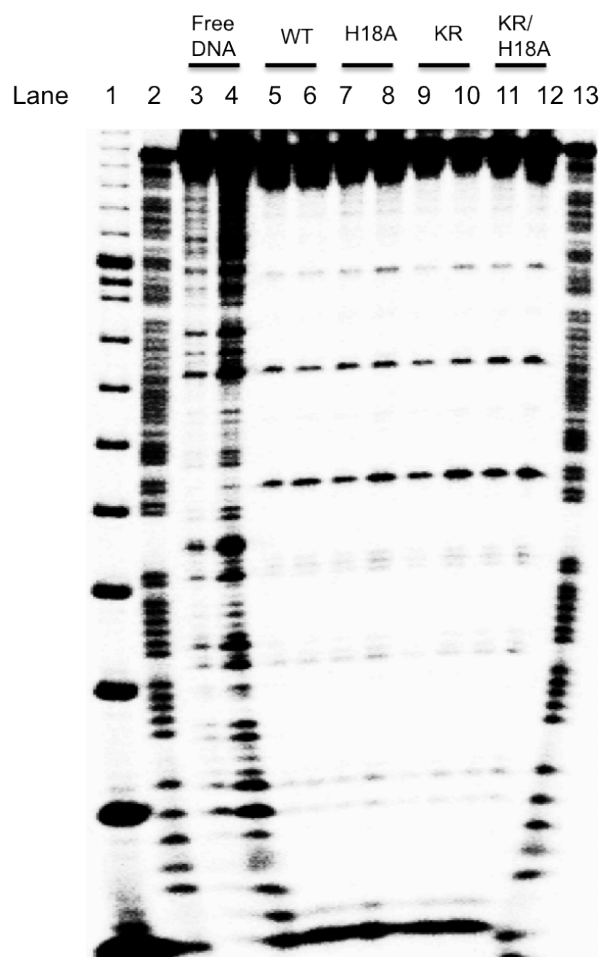
(C)



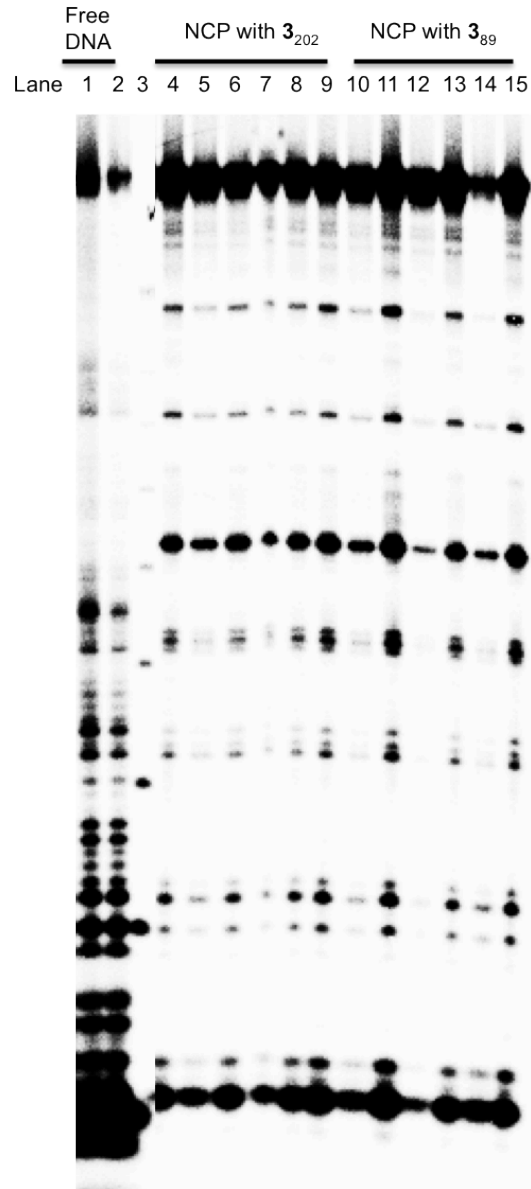
**Figure S9.** Chromatograms of reacting **4** with (A) protected lysine (**6a**); (B) protected histidine (**6b**); (C) protected cysteine (**6c**).



**Figure S10.** Representative chromatogram for purifying oligonucleotides containing **3** (oligonucleotide S31).

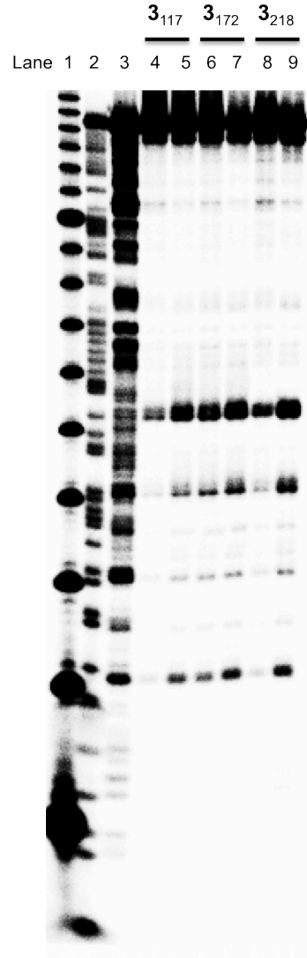


**Figure S11.** DNase I footprinting of NCPs containing  $3_{89}$  and  $3_{205}$ . Lane 1: DNA 10-bp ladder; lane 2: A + G sequencing on DNA duplex containing  $3_{89}$ ; lane 3-4: DNase I digested free 601 DNA containing  $3_{89}$  and  $3_{205}$ ; lane 5: DNase I digested NCPs composed of WT histone octamers with  $3_{205}$ ; lane 6: DNase I digested NCPs composed of WT histone octamers with  $3_{89}$ ; lane 7-12: DNase I digested NCPs composed of histone H4 H18A variant (lane 7-8) or H4 K5, 8, 12, 16, 20R (lane 9-10) or H4 K5, 8, 12, 16, 20R/H18A (lane 11-12) with  $3_{89}$ ; lane 13: A + G sequencing on DNA duplex containing  $3_{205}$ . DNase I was added to each sample in the following amounts: 0.1 unit for lanes 7, 9, 11; 0.2 unit for lanes 5, 6, 8, 10, 12.

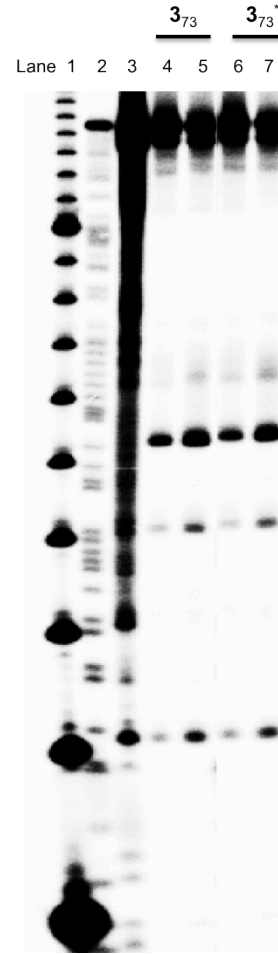


**Figure S12.** DNase I footprinting of NCPs containing  $3_{89}$  and  $3_{202}$ . Lane 1-2: DNase I digested free 601 DNA containing  $3_{89}$  and  $3_{202}$ ; lane 3: DNA 10-bp ladder; lane 4-9: DNase I digested NCPs composed of WT histone proteins (lane 4-5) or histone H4 K16, 20A variant (lane 6-7) or histone H4 K5, 8, 12A variant (lane 8-9) with  $3_{202}$ ; lane 10-15: DNase I digested NCPs composed of histone H4 K16, 20A variant (lane 10-11) or histone H4 K5, 8, 12A variant (lane 12-13) or histone H4 del 1-20 (lane 14-15) with  $3_{89}$ ; DNase I was added to each sample in the following amounts: 0.1 unit for lanes 4, 6, 8, 10, 12, 14; 0.2 unit for lanes 5, 7, 9, 11, 13, 15.

(A)

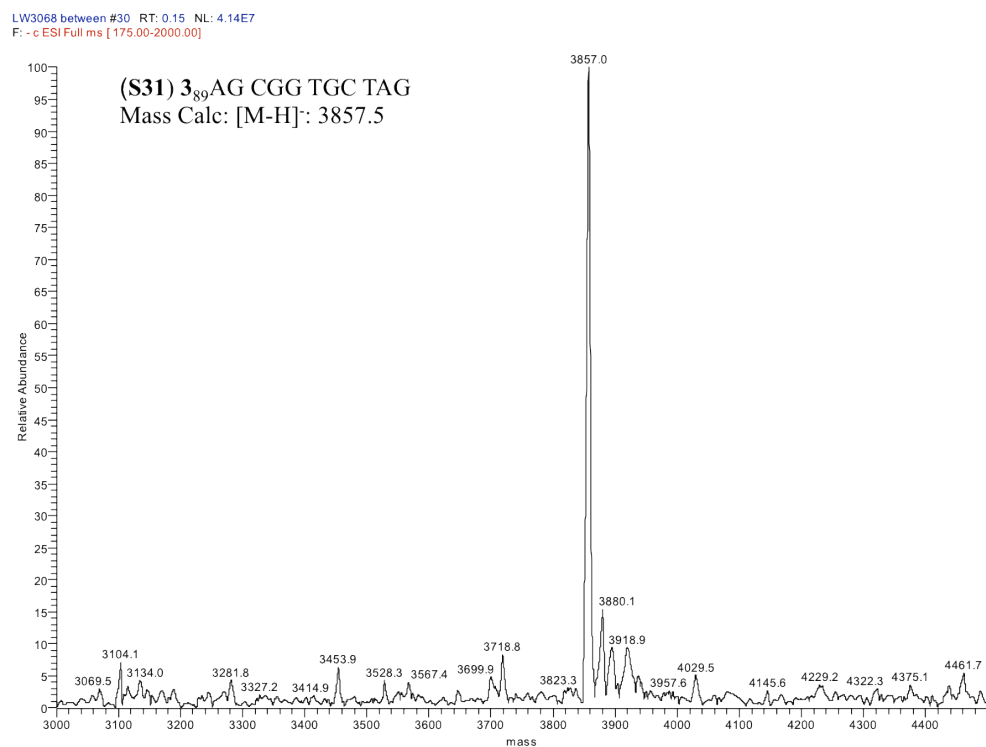
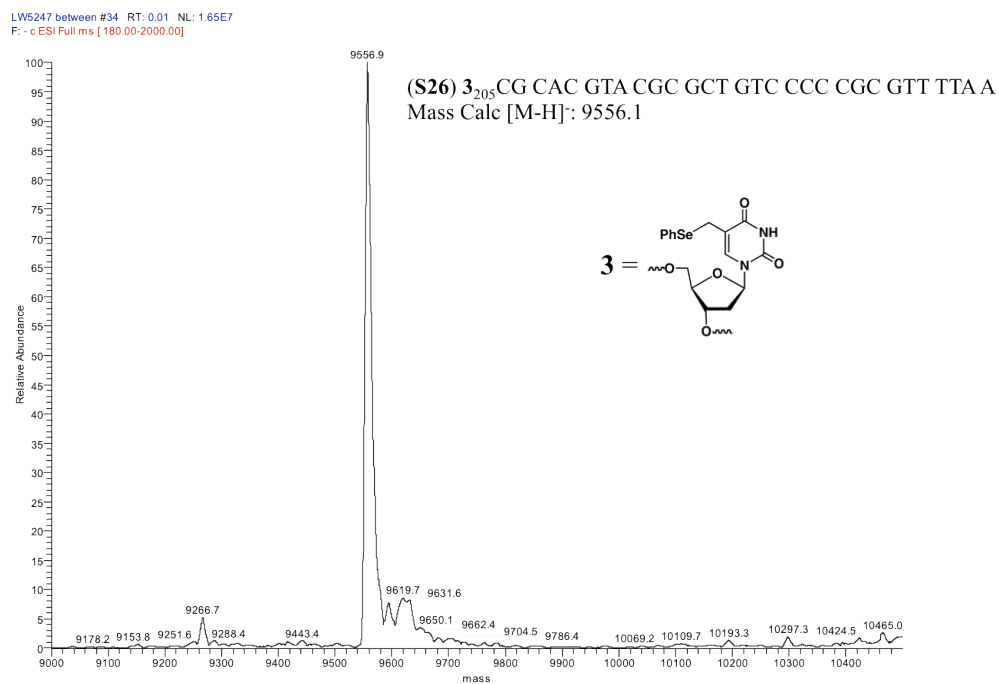


(B)

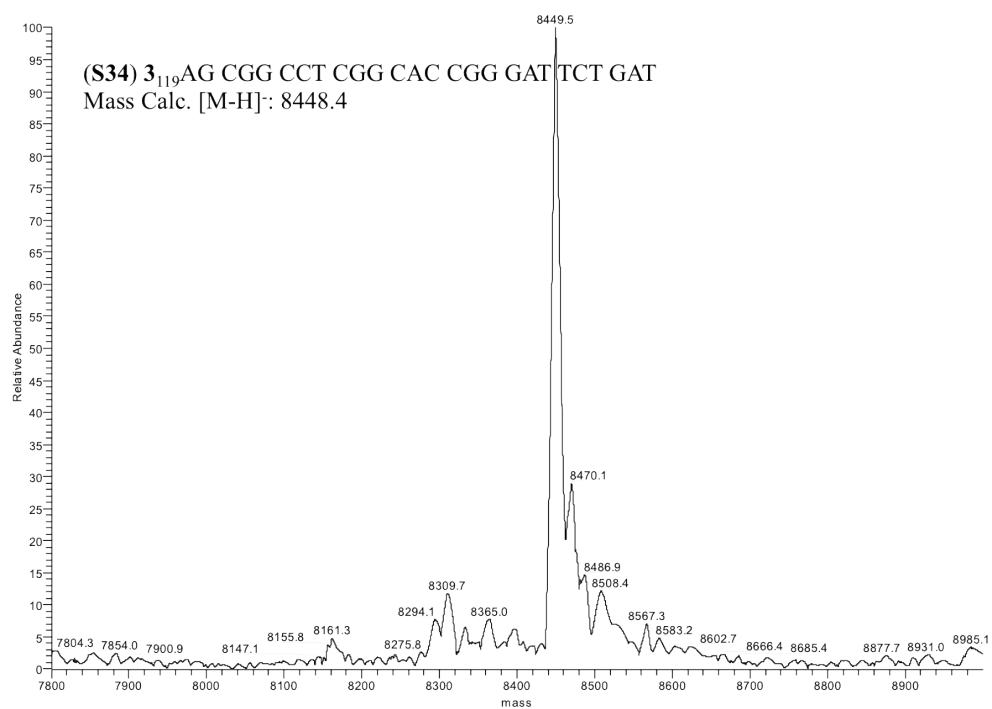


**Figure S13.** (A) DNase I footprinting of NCPs containing  $3_{117}$  (lanes 4-5),  $3_{172}$  (lanes 6-7) and  $3_{218}$  (lanes 8-9). Lane 1: DNA 10-bp ladder; lane 2: A + G sequencing on DNA duplex containing  $3_{117}$ ; lane 3: DNase I digested free 601 DNA containing  $3_{117}$ ; DNase I was added to each sample in the following amounts: 0.1 unit for lanes 4, 6, 8; 0.2 unit for lanes 5, 7, 9. (B) DNase I footprinting of NCPs containing  $3_{73}$  (lanes 4-5) and  $3_{73}^*$  (lanes 6-7, \* indicates T3A sequence); lane 1: DNA 10-bp ladder; lane 2: A + G sequencing on DNA duplex containing  $3_{73}$ ; lane 3: DNase I digested free 601 DNA containing  $3_{73}$ ; DNase I was added to each sample in the following amounts: 0.1 unit for lanes 4, 6; 0.2 unit for lanes 5, 7.

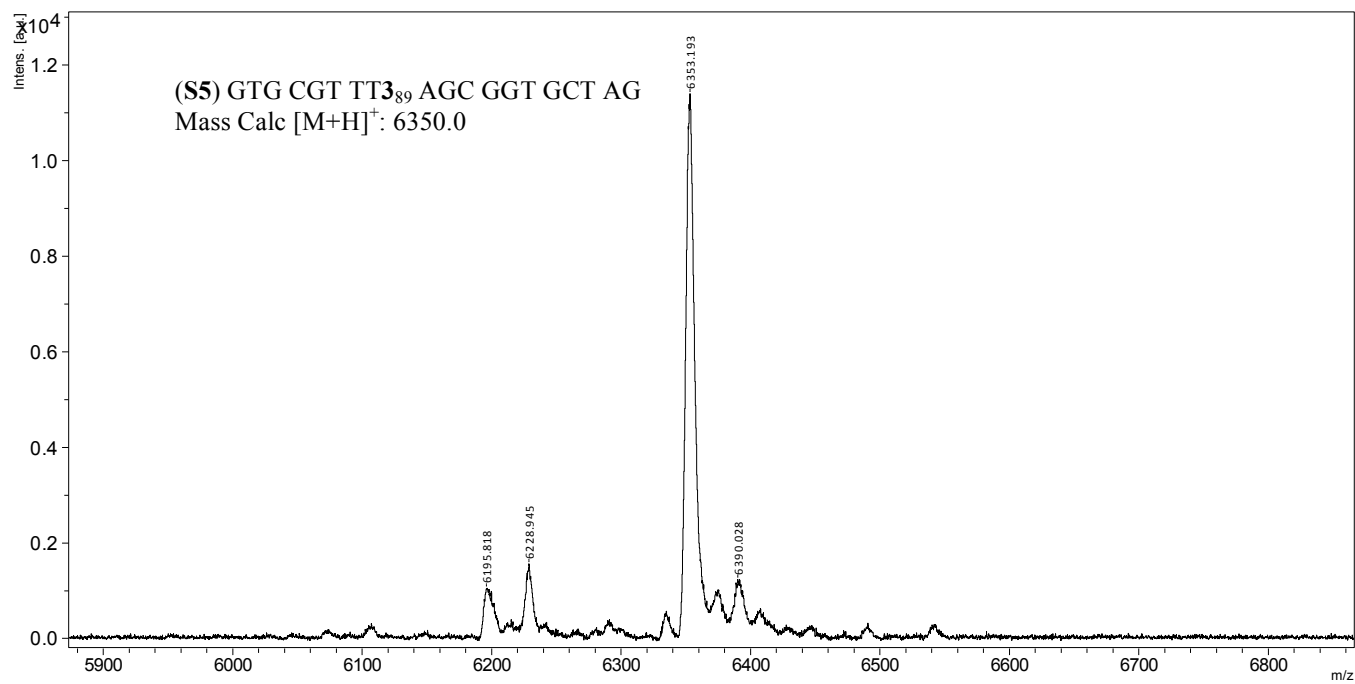
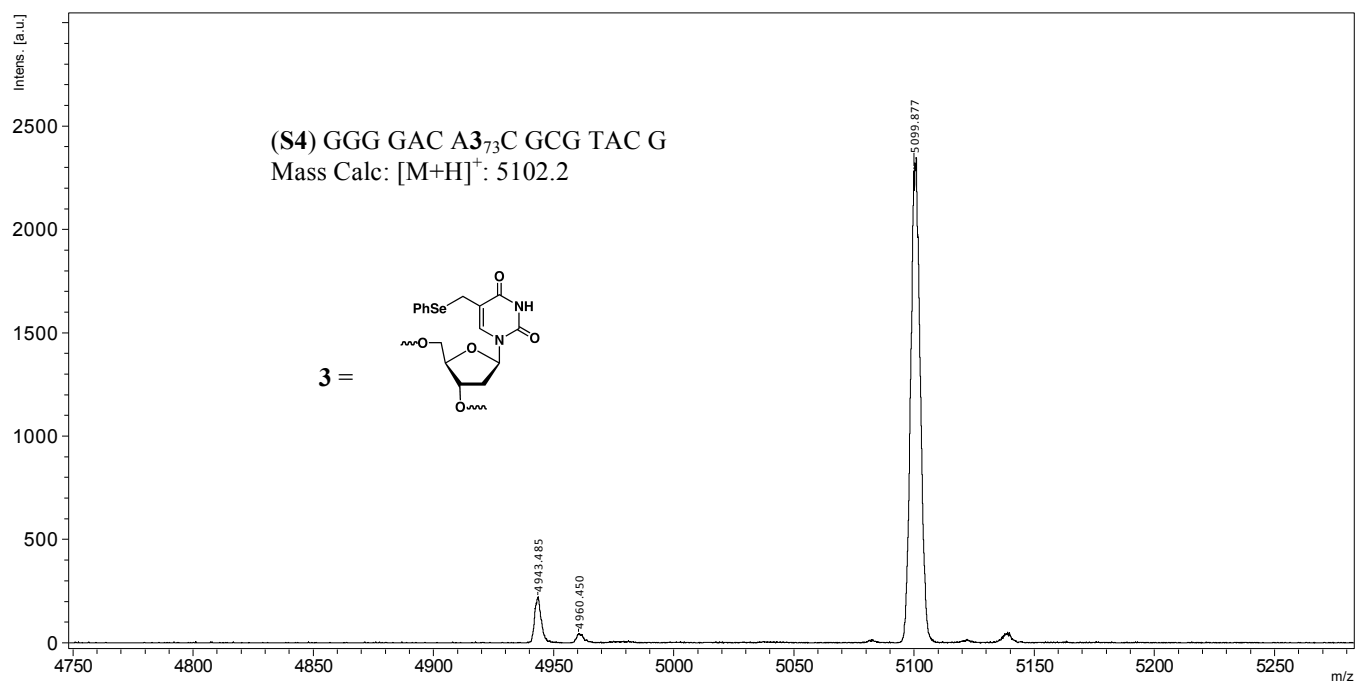
**Figure S14.** ESI-MS spectra of modified oligonucleotides.

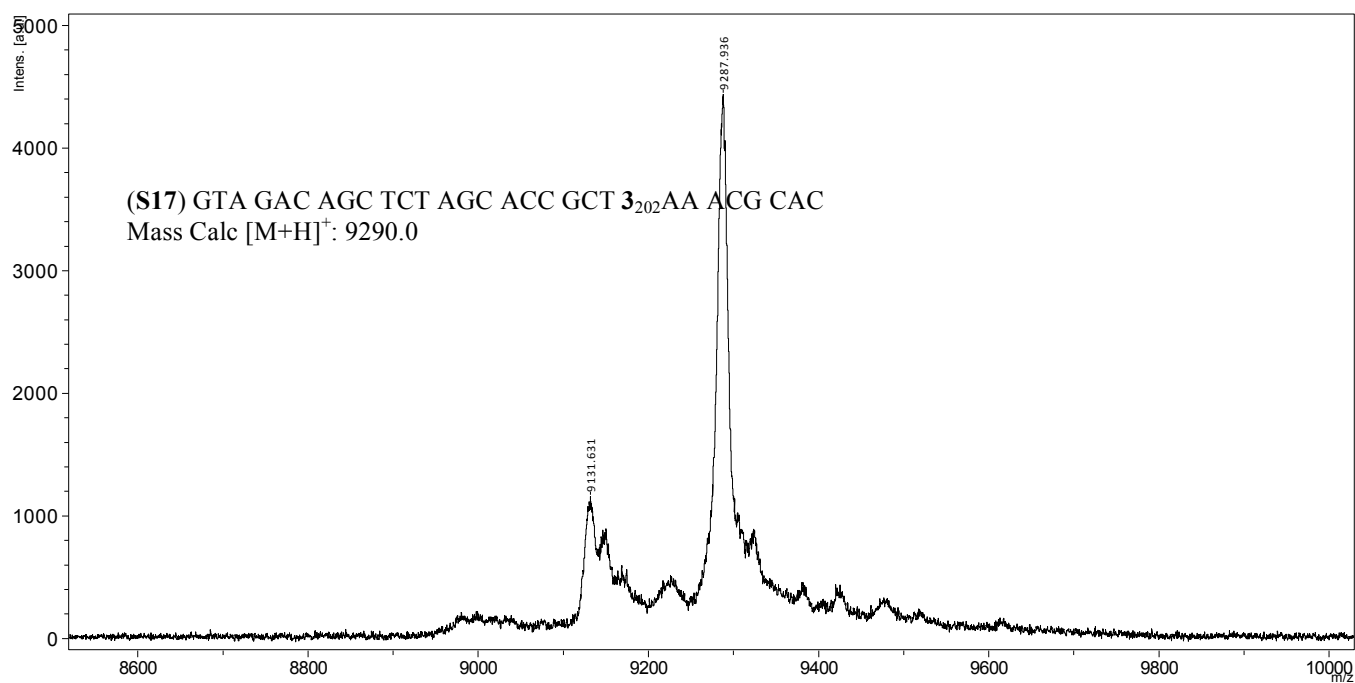
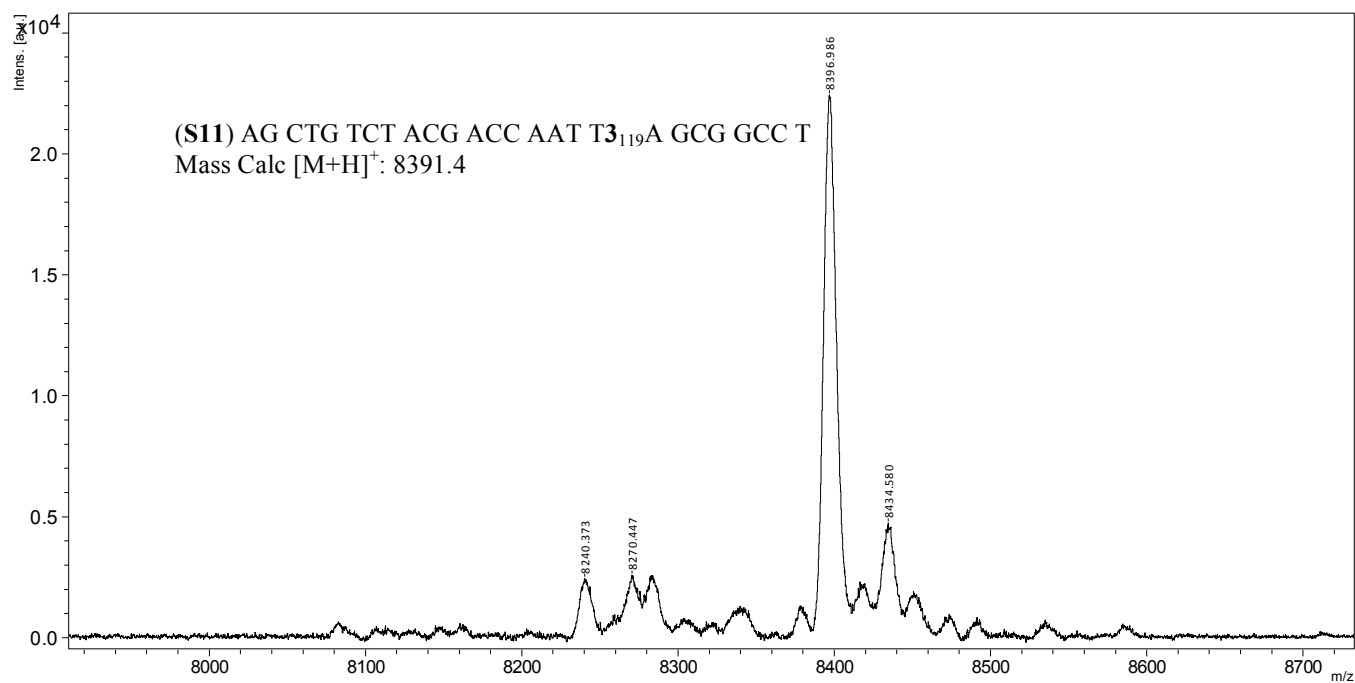


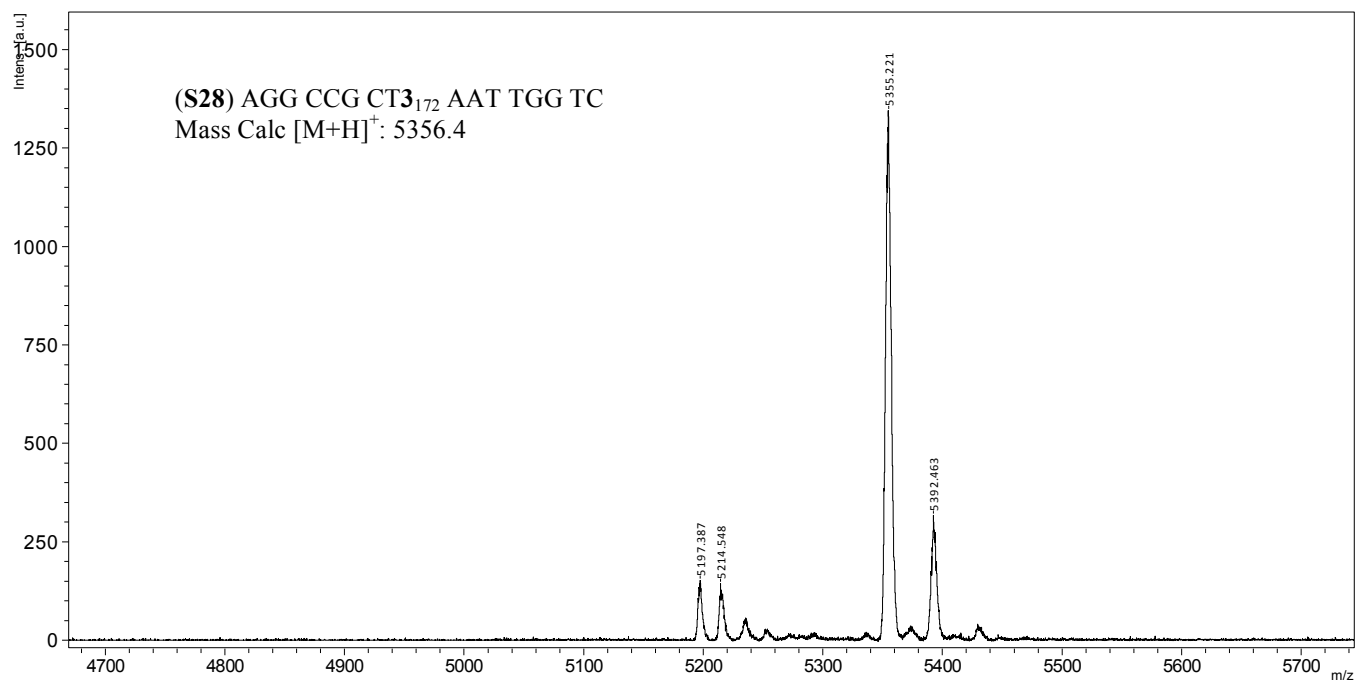
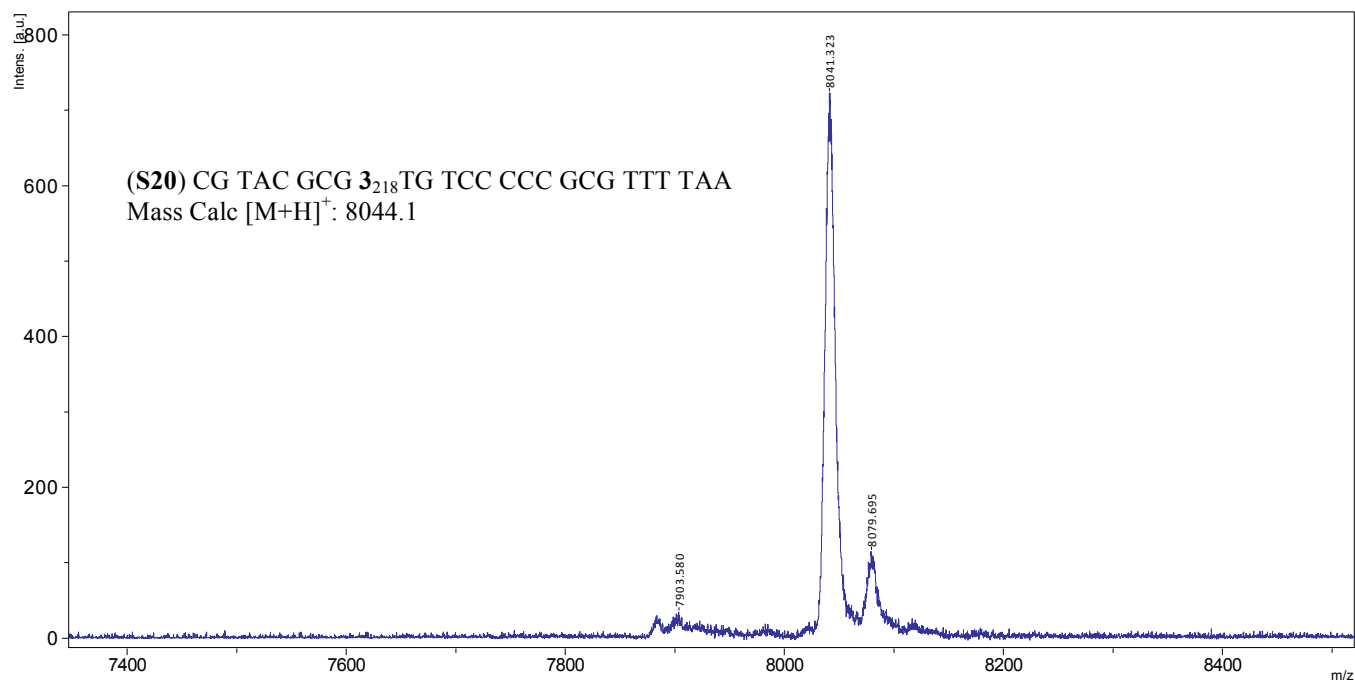
LW3067 between #30 RT: 0.01 NL: 1.20E8  
F: -c ESI Full ms [175.00-2000.00]

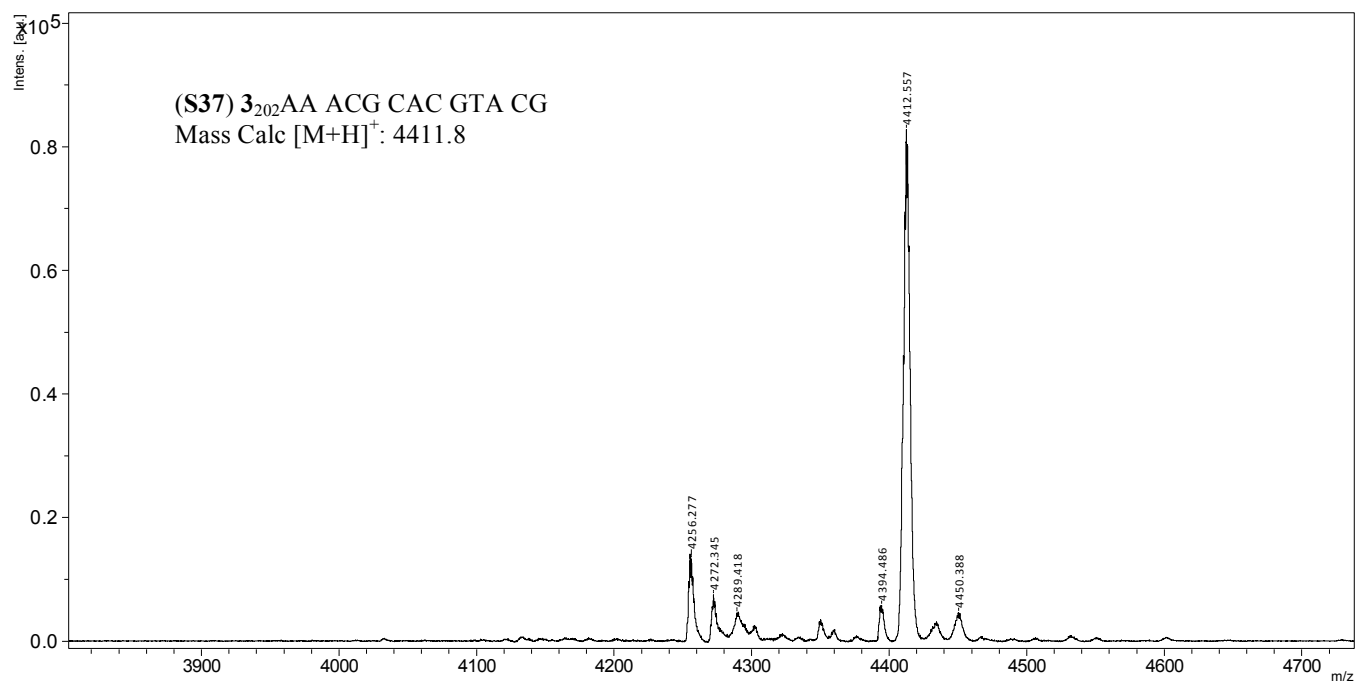
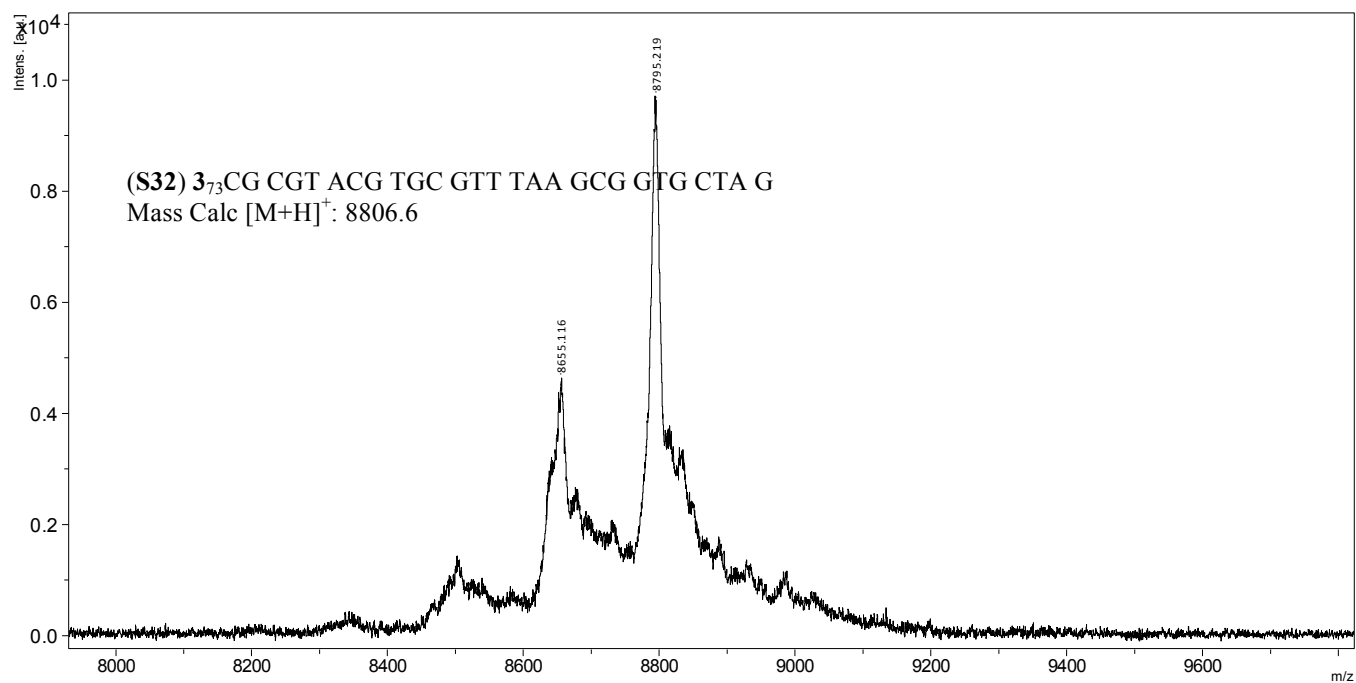


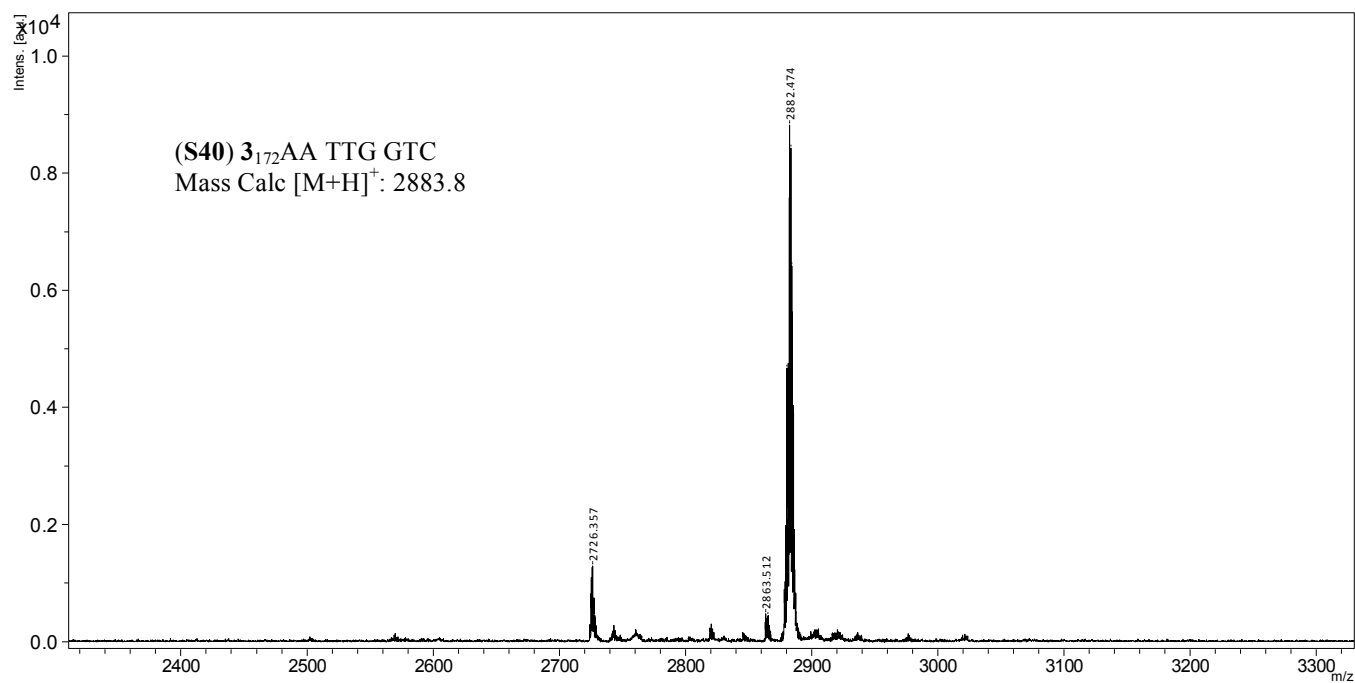
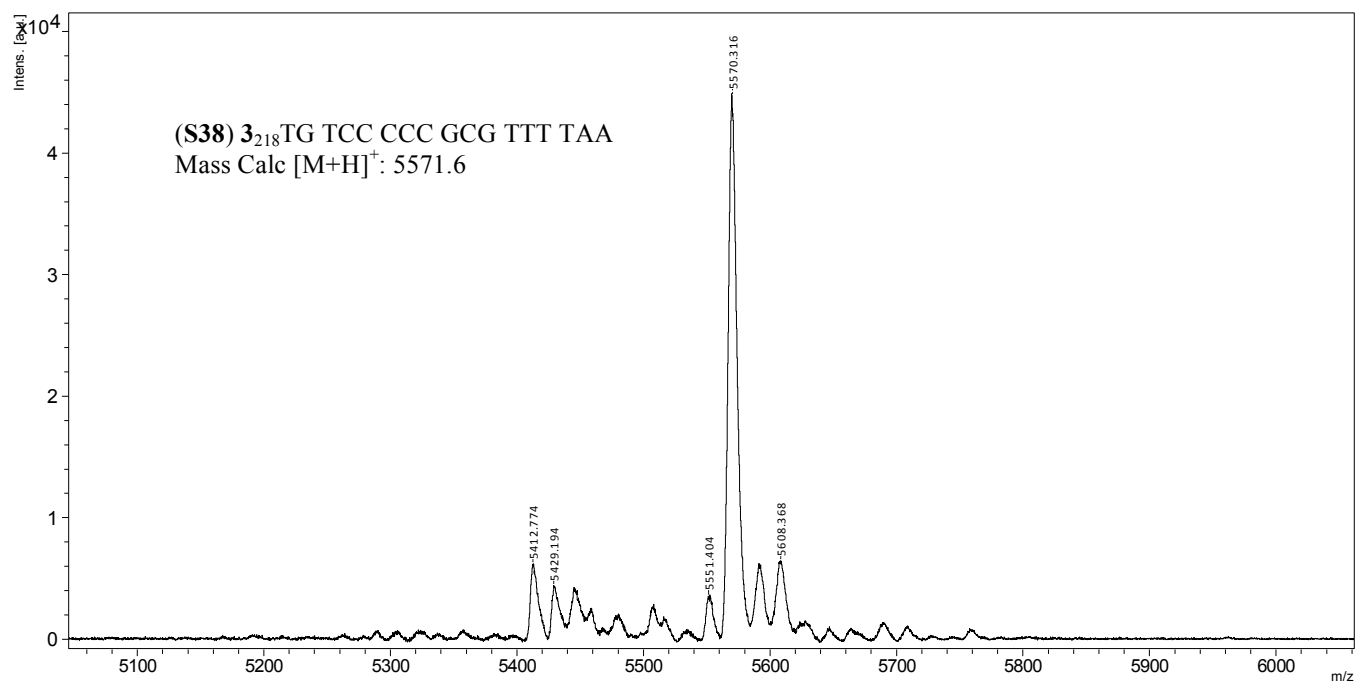
**Figure S15.** MALDI-TOF mass spectra of modified oligonucleotides (the laser of MALDI-TOF instrument can cleave the phenyl selenide group, resulting in a peak with 156 Da less than the expected molecular weight).



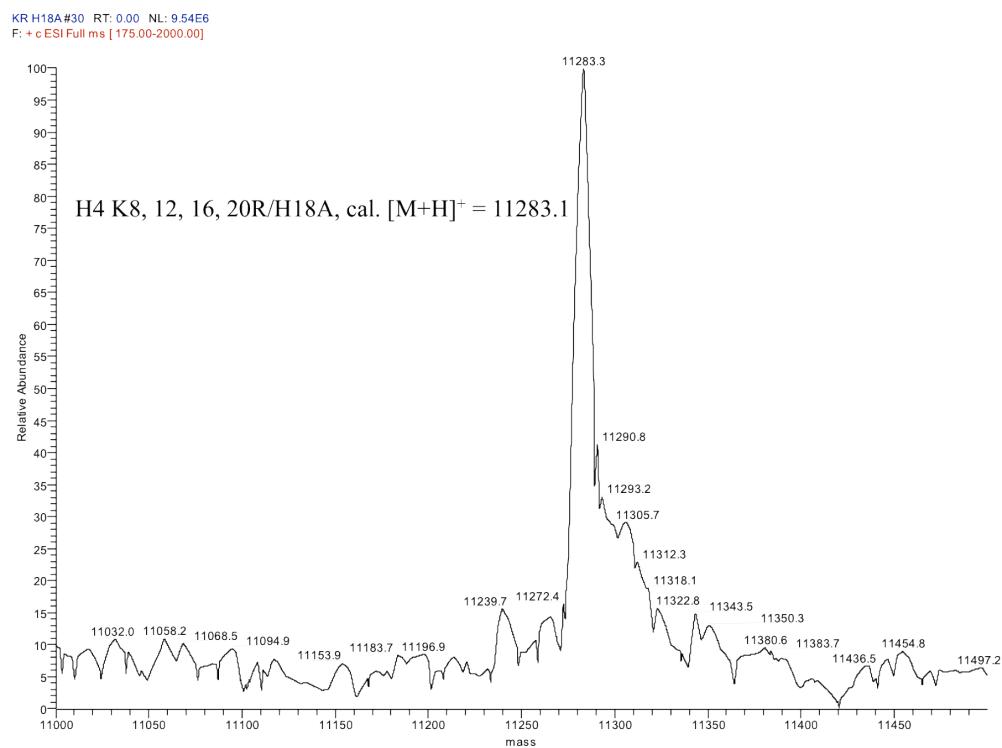
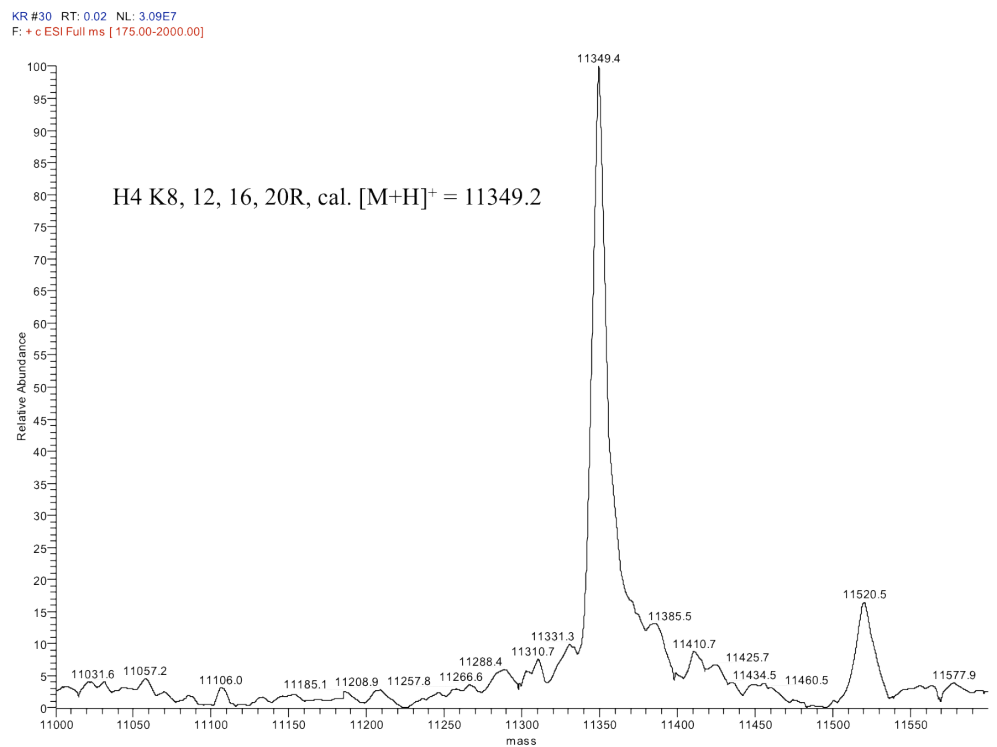


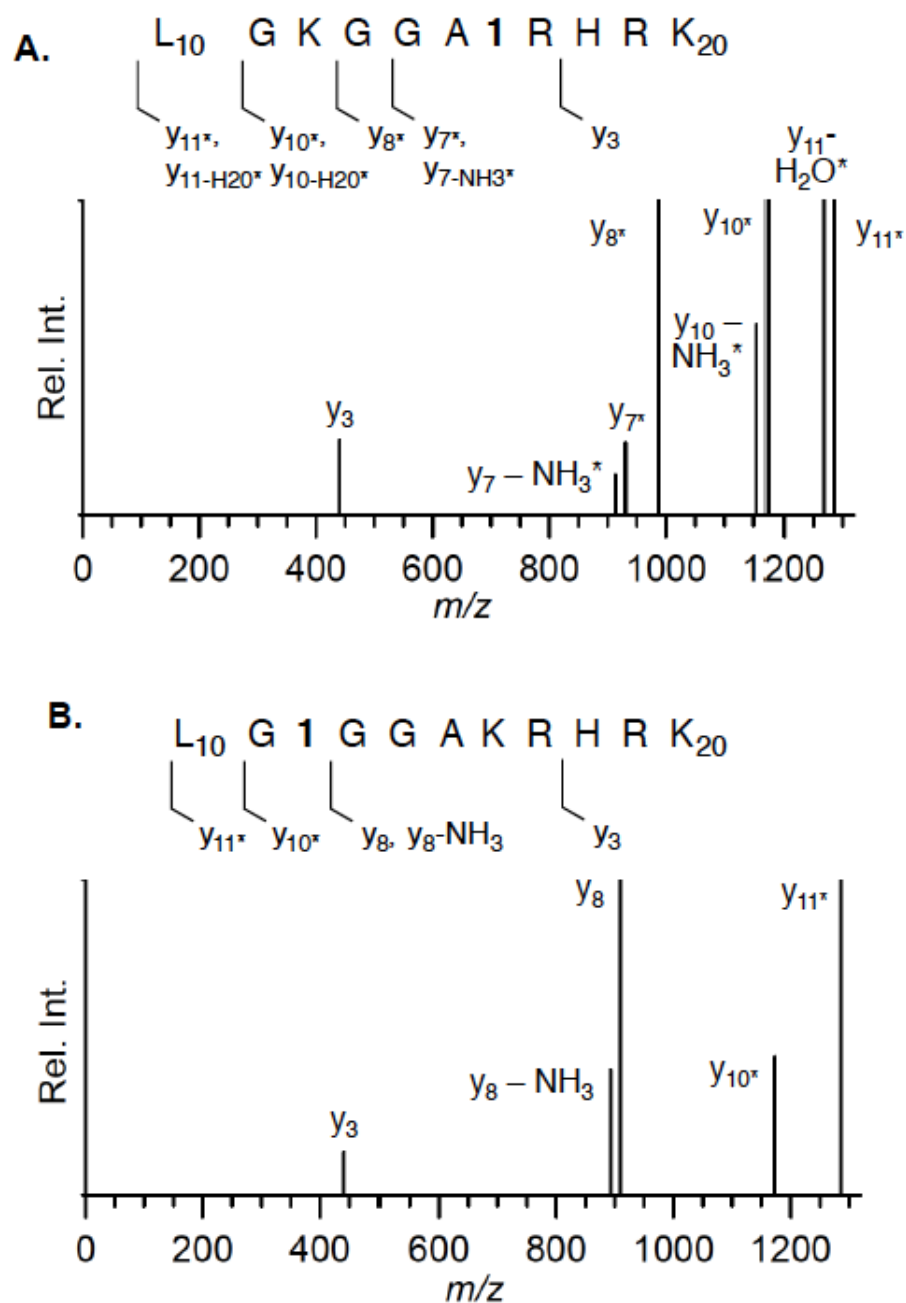






**Figure S16.** ESI mass spectra of histone H4 K8, 12, 16, 20R and K8, 12, 16, 20R/H18A variants.





**Figure S17.** LC-MS/MS spectra of peptide fragments following thermolysin digestion of histone H4. A. Fragmentation pattern of peptide containing **1** at Lys16 (K16). B. Fragmentation pattern of peptide containing **1** at Lys12 (K12).

**Table S1.** Calculated and observed fragmentation patterns for peptides in Figures 2, 3, and Figure S17.

	<b>Cal.</b>	<b>Found</b>
<b>Figure 2A.</b> Fragment 20-23 (K20 mod.)		
y1	175.1195	175.1200
y2	288.2036	288.2056
y3	387.2720	387.2746
a2*	278.1869	278.1870
y4*	593.3775	593.3762
y2-NH3	271.1770	271.1766

**Figure 2B.** Fragment 4-17 (K16 mod.)

b2*	228.1348	228.1384
b3*	285.1563	285.1574
y1	175.1195	175.1212
y4*	509.2836	509.2855
y5*	566.3051	566.3029
y6*	736.4106	736.4048
y7*	793.4321	793.4268
y9*	963.5376	963.5257
y12*	1247.6862	1247.6775
y14*	1474.8131	1474.8136

**Figure 2C.** Fragment 4-17 (K12 mod.)

b8*	739.4103	739.4103
y1	175.1195	175.1222
y5*	530.3051	530.3022
y14*	1474.8131	1474.8159

	Cal.	Found
--	------	-------

**Figure 3.** Fragment 1-9 (K8 mod.)

b5	486.2788	486.2782
b7	600.3218	600.3230
a7	572.3268	572.3288
a8*	778.4324	778.4362
y9*	881.4594	881.4558

**Figure S17A.** Fragment 10-20 (K16 mod.)

y3	440.2734	440.2713
y7*	930.5386	930.5446
y8*	987.5601	987.5624
y10*	1172.6765	1172.6753
y10-H2O*	1154.6659	1154.6577
y11-H2O*	1267.7500	1267.7426
y11*	1285.7605	1285.7614
y7-NH3*	913.5121	913.5135

**Figure S17B.** Fragment 10-20 (K12 mod.)

y3	440.2734	440.2760
y8	909.5495	909.5525
y10*	1172.6765	1172.6761
y11*	1285.7605	1285.7588
y8-NH3	892.5229	892.5279